



OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/276,935C

DATE: 12/12/2002
TIME: 16:44:08

Errors on pp.
3-5

Input Set : A:\PU3474US 11-02 Seqlist.txt
Output Set: N:\CRF4\12122002\I276935C.raw

4 <110> APPLICANT: KLEWER, Steven A.
5 JONES, Stacey A.
6 WILLSON, Timothy M.
8 <120> TITLE OF INVENTION: AN ORPHAN NUCLEAR RECEPTOR
11 <130> FILE REFERENCE: PU3474US2
13 <140> CURRENT APPLICATION NUMBER: 09/276,935C
C--> 14 <141> CURRENT FILING DATE: 2002-11-27
16 <150> PRIOR APPLICATION NUMBER: 60/079,593
17 <151> PRIOR FILING DATE: 1998-03-27
19 <160> NUMBER OF SEQ ID NOS: 18
21 <170> SOFTWARE: FastSEQ for Windows Version 4.0
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 20
25 <212> TYPE: DNA
26 <213> ORGANISM: Artificial Sequence
28 <220> FEATURE:
29 <223> OTHER INFORMATION: Probe
31 <400> SEQUENCE: 1
32 ctqctqcqca tccaggacat 20
34 <210> SEQ ID NO: 2
35 <211> LENGTH: 45
36 <212> TYPE: DNA
37 <213> ORGANISM: Artificial Sequence
39 <220> FEATURE:
40 <223> OTHER INFORMATION: Probe
42 <400> SEQUENCE: 2
43 gggtgtgggg aatccaccac catggaggtg agacccaaag aaage 45
45 <210> SEQ ID NO: 3
46 <211> LENGTH: 34
47 <212> TYPE: DNA
48 <213> ORGANISM: Artificial Sequence
50 <220> FEATURE:
51 <223> OTHER INFORMATION: Probe
53 <400> SEQUENCE: 3
54 gggtgtgggg gatccctcagc tacctgtgtat gccg 34
56 <210> SEQ ID NO: 4
57 <211> LENGTH: 31
58 <212> TYPE: DNA
59 <213> ORGANISM: Artificial Sequence
61 <220> FEATURE:
62 <223> OTHER INFORMATION: Probe
64 <400> SEQUENCE: 4
65 gatcagacaaq ttcatgtaaq: tcatctatat c 31

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/276,935C

DATE: 12/12/2002
TIME: 16:44:08

Input Set : A:\PU3474US 11-02 Seqlist.txt
Output Set: N:\CRF4\12122002\I276935C.raw

67 <210> SEQ ID NO: 5
68 <211> LENGTH: 29
69 <212> TYPE: DNA
70 <213> ORGANISM: Artificial Sequence
72 <220> FEATURE:
73 <223> OTHER INFORMATION: Probe
75 <400> SEQUENCE: 5
76 gatcaatatg aactcaaagg aggtcagtg 29
78 <210> SEQ ID NO: 6
79 <211> LENGTH: 29
80 <212> TYPE: DNA
81 <213> ORGANISM: Artificial Sequence
83 <220> FEATURE:
84 <223> OTHER INFORMATION: Probe
86 <400> SEQUENCE: 6
87 gatcaatatg aactcaaagg aggtcagtg 29
89 <210> SEQ ID NO: 7
90 <211> LENGTH: 29
91 <212> TYPE: DNA
92 <213> ORGANISM: Artificial Sequence
94 <220> FEATURE:
95 <223> OTHER INFORMATION: Probe
97 <400> SEQUENCE: 7
98 gatcaatatg ttctcaaagg agaacagtg 29
100 <210> SEQ ID NO: 8
101 <211> LENGTH: 29
102 <212> TYPE: DNA
103 <213> ORGANISM: Artificial Sequence
105 <220> FEATURE:
106 <223> OTHER INFORMATION: Probe
108 <400> SEQUENCE: 8
109 gatcaataac aactcaaagg aggtcagtg 29
111 <210> SEQ ID NO: 9
112 <211> LENGTH: 32
113 <212> TYPE: DNA
114 <213> ORGANISM: Artificial Sequence
116 <220> FEATURE:
117 <223> OTHER INFORMATION: Probe
119 <400> SEQUENCE: 9
120 gatgcagaca gttcatgaag ttcatctaga tc 32
122 <210> SEQ ID NO: 10
123 <211> LENGTH: 11
124 <212> TYPE: PRT
125 <213> ORGANISM: Artificial Sequence
127 <220> FEATURE:
128 <223> OTHER INFORMATION: Probe
130 <400> SEQUENCE: 10
131 Met Lys Lys Gly His His His His His Gly
132 1 5 10

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/276,935C

DATE: 12/12/2002

TIME: 16:44:08

Input Set : A:\PU3474US 11-02 Seqlist.txt
Output Set: N:\CRF4\12122002\I276935C.raw

135 <210> SEQ ID NO: 11
 136 <211> LENGTH: 316
 137 <212> TYPE: PRT
 138 <213> ORGANISM: Artificial Sequence
 140 <220> FEATURE:
 141 <223> OTHER INFORMATION: Protein - *must explain genetic source -
See cover summary sheet item 11*
 143 <400> SEQUENCE: 11
 144 Met Lys Lys Gly His His His His His Gly Ser Glu Arg Thr Gly
 145 1 5 10 15
 146 Thr Gln Pro Leu Gly Val Gln Gly Leu Thr Glu Glu Gln Arg Met Met
 147 20 25 30
 148 Ile Arg Glu Leu Met Asp Ala Gln Met Lys Thr Phe Asp Thr Thr Phe
 149 35 40 45
 150 Ser His Phe Lys Asn Phe Arg Leu Pro Gly Val Leu Ser Ser Gly Cys
 151 50 55 60
 152 Glu Leu Pro Glu Ser Leu Gln Ala Pro Ser Arg Glu Glu Ala Ala Lys
 153 65 70 75 80
 154 Trp Ser Gln Val Arg Lys Asp Leu Cys Ser Leu Lys Val Ser Leu Gln
 155 85 90 95
 156 Leu Arg Gly Glu Asp Gly Ser Val Trp Asn Tyr Lys Pro Pro Ala Asp
 157 100 105 110
 158 Ser Gly Gly Lys Glu Ile Phe Ser Leu Leu Pro His Met Ala Asp Met
 159 115 120 125
 160 Ser Thr Tyr Met Phe Lys Gly Ile Ile Ser Phe Ala Lys Val Ile Ser
 161 130 135 140
 162 Tyr Phe Arg Asp Leu Pro Ile Glu Asp Gln Ile Ser Leu Leu Lys Gly
 163 145 150 155 160
 164 Ala Ala Phe Glu Leu Cys Gln Leu Arg Phe Asn Thr Val Phe Asn Ala
 165 165 170 175
 166 Glu Thr Gly Thr Trp Glu Cys Gly Arg Leu Ser Tyr Cys Leu Glu Asp
 167 180 185 190
 168 Thr Ala Gly Gly Phe Gln Gln Leu Leu Glu Pro Met Leu Lys Phe
 169 195 200 205
 170 His Tyr Met Leu Lys Lys Leu Gln Leu His Glu Glu Tyr Val Leu
 171 210 215 220
 172 Met Gln Ala Ile Ser Leu Phe Ser Pro Asp Arg Pro Gly Val Leu Gln
 173 225 230 235 240
 174 His Arg Val Val Asp Gln Leu Gln Glu Gln Phe Ala Ile Thr Leu Lys
 175 245 250 255
 176 Ser Tyr Ile Glu Cys Asn Arg Pro Gln Pro Ala His Arg Phe Leu Phe
 177 260 265 270
 178 Leu Lys Ile Met Ala Met Leu Thr Glu Leu Arg Ser Ile Asn Ala Gln
 179 275 280 285
 180 His Thr Gln Arg Leu Leu Arg Ile Gln Asp Ile His Pro Phe Ala Thr
 181 290 295 300
 182 Pro Leu Met Gln Glu Leu Phe Gly Ile Thr Gly Ser
 183 305 310 315
 186 <210> SEQ ID NO: 12
 187 <211> LENGTH: 242

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/276,935C

DATE: 12/12/2002

TIME: 16:44:08

Input Set : A:\PU3474US 11-02 Seqlist.txt
 Output Set: N:\CRF4\12122002\I276935C.raw

188 <212> TYPE: PRT
 189 <213> ORGANISM: Artificial Sequence
 191 <220> FEATURE:
 192 <223> OTHER INFORMATION: Protein *Some error*
 194 <400> SEQUENCE: 12
 195 Met Lys Lys Gly Ser Ala Asn Glu Asp Met Pro Val Glu Arg Ile Leu
 196 1 5 10 15
 197 Glu Ala Glu Leu Ala Val Glu Pro Lys Thr Glu Thr Tyr Val Glu Ala
 198 20 25 30
 199 Asn Met Gly Leu Asn Pro Ser Ser Pro Asn Asp Pro Val Thr Asn Ile
 200 35 40 45
 201 Cys Gln Ala Ala Asp Lys Gln Leu Phe Thr Leu Val Glu Trp Ala Lys
 202 50 55 60
 203 Arg Ile Pro His Phe Ser Glu Leu Pro Leu Asp Asp Gln Val Ile Leu
 204 65 70 75 80
 205 Leu Arg Ala Gly Trp Asn Glu Leu Leu Ile Ala Ser Phe Ser His Arg
 206 85 90 95
 207 Ser Ile Ala Val Lys Asp Gly Ile Leu Leu Ala Thr Gly Leu His Val
 208 100 105 110
 209 His Arg Asn Ser Ala His Ser Ala Gly Val Gly Ala Ile Phe Asp Arg
 210 115 120 125
 211 Val Leu Thr Glu Leu Val Ser Lys Met Arg Asp Met Gln Met Asp Lys
 212 130 135 140
 213 Thr Glu Leu Gly Cys Leu Arg Ala Ile Val Leu Phe Asn Pro Asp Ser
 214 145 150 155 160
 215 Lys Gly Leu Ser Asn Pro Ala Glu Val Glu Ala Leu Arg Glu Lys Val
 216 165 170 175
 217 Tyr Ala Ser Leu Glu Ala Tyr Cys Lys His Lys Tyr Pro Glu Gln Pro
 218 180 185 190
 219 Gly Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ala Leu Arg Ser Ile
 220 195 200 205
 221 Gly Leu Lys Cys Leu Glu His Leu Phe Phe Lys Leu Ile Gly Asp
 222 210 215 220
 223 Thr Pro Ile Asp Thr Phe Leu Met Glu Met Leu Glu Ala Pro His Gln
 224 225 230 235 240
 225 Met Thr
 229 <210> SEQ ID NO: 13
 230 <211> LENGTH: 2146
 231 <212> TYPE: DNA
 232 <213> ORGANISM: Artificial Sequence
 234 <220> FEATURE:
 235 <223> OTHER INFORMATION: Probe
 237 <400> SEQUENCE: 13
 238 tggaaatataq tgagagaca agatgtctc atatccgggg aaatcataac ctatgactag 60
 239 gacgggaaga ggaaggactq ctttacttc agtgggaatc tcggcccaq cctgeaagcc 120
 240 aagtqitcac aqgagaaaa gcaagagaat aqctaatac tcctgtctg aacaaggcag 180
 241 cqgcttcctg qtaaaqctac tccttgatcg atcctttgca ccggatttgtt caaagtggac 240
 242 cccaggggag aagtccgggc aaqaaactta ccaccaagca gtcacaaggg cccagaagca 300
 243 aaccttqqaqq tgagacccaa agaaaagctgg aaccatgtq actttgtaca ctgtgaggac 360

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/276,935C

DATE: 12/12/2002

TIME: 16:44:08

Input Set : A:\PU3474US 11-02 Seqlist.txt

Output Set: N:\CRF4\12122002\I276935C.raw

244 acaqagtctg ttectggaaa gcccagtg aacgcaqatg aggaagtctgg aggccccaa 420
 245 aictgcctgtq tatgtggggaa caagggccact qgctatcaact tcaatgtcat gacatgtqaa 480
 246 qqatqcaaqg qcttttcag gaggggccatg aaaaqgeaaccg cccggctgag gtggcccttc 540
 247 cggaaaggccq cctgcgagat caccggqaag accccqgcac agtgcgcaggc ctggccctg 600
 248 cgcagaatgtcc tggagagccqg catgaagaqaa qagatqatca tgccqacgaa ggcctggag 660
 249 gagagggggg ccttqatcaa gggaaqaaa agtgaacgga caggactca gccactgggaa 720
 250 gtgcaggggc tgacagaaqqa qcagcggatg atqatcagg agtqatggaa cgctcagatg 780
 251 aaaacctttt acactacattt ctccatttc aqauatttc ggttgcagg ggtqcttagc 840
 252 aqgtggctcg agtggccaga gtctctgcag gcccatacgaa qggaagaagc tgccaaagtgg 900
 253 agccagggtcc gggaaagatct gtgtctttt aqgtqcttc tgccagtcg qggggaggat 960
 254 ggcagtttgc ggaactacaa acccccaccc qacagtggcg gggaaagatg ctctccctt 1020
 255 ctggccccaca tggctgacat gtcacacccat atgttcaaaq gcatcatcg ctgtggccaaa 1080
 256 gtcatcttactt acttcaggga ctgtccccatc gaggaccaga tctccctgtc qaaaggggcc 1140
 257 gettgcgac tgggtcaact gagatcaac acagtgttca acgcggagac tggaaacctgg 1200
 258 gagtggtggcc ggctgttca ctgttggaa qacactgcag gtggcttcca gcaacttctt 1260
 259 ctggagccca tggtaaaattt ccactacatg ctgaaagaage tgccagtcg tgaggaggag 1320
 260 tatgtgttca tgcaggccat ctccctttt tccccagacc gcccagggtgt gctgcac 1380
 261 cgcgtgggtgg accagctgca ggagcaattt gecattactt tqaqgttcaatgaaatgc 1440
 262 aatccggcccc accctgttca taggttcttg ttccttggaa tcatggctat qctcaccggq 1500
 263 ctccgcagca tcaatgttca gacacacccag cggctqctgc gcatccaggaa catacacc 1560
 264 ttgtctacgc ccctcatgca ggagtgttc gycatcacaq gttagctgac qgctggccctt 1620
 265 ggtgtacacc tccgagggcc acccagaccc agagccctt gacccggccac tcccggccaa 1680
 266 agacagatgg acactgcca gggccacaa tgccctgtc gctgtctcc ctagggaaattt 1740
 267 cctgctatgtc cagctggctca gcatccttca ggaaggacat gggggccccca caccggcagt 1800
 268 tcaagtctgttca gggagtgttca gggaggacatccacatc ttacgtggag agtgcacttca cctgttaggtc 1860
 269 aggaccatca gagaggcaag gttggccctt cctttttttt ggcctgttgg tctggggaga 1920
 270 aatccctcaatc atccactaa agtgcacttca agtgcacttca tggtaaaatggg accaagcgac caaggatagg 1980
 271 ccatctgggg tctatggccaa catacccaacq ttgttgcgt tccctgagtct tttcattgt 2040
 272 acctcttataata gtcctgttca ccacttccca ctcgttcccc tccctttccq agtgcattttt 2100
 273 tgggctccag gctgtacttccatc atcggcaggat gcatgactat ctgtgg 2146

275 <210> SEQ ID NO: 14

276 <211> LENGTH: 434

277 <212> TYPE: PRT

278 <213> ORGANISM: Artificial Sequence

280 <220> FEATURE:

281 <223> OTHER INFORMATION, *Protein* -- *Some t* *not*

283 <400> SEQUENCE: 14

284 Leu Glu Val Arg Pro Lys Glu Ser Trp Asn His Ala Asp Phe Val His
 285 1 5 10 15
 286 Cys Glu Asp Thr Glu Ser Val Pro Gly Lys Pro Ser Val Asn Ala Asp
 287 20 25 30
 288 Glu Glu Val Gly Gly Pro Gln Ile Cys Arg Val Cys Gly Asp Lys Ala
 289 35 40 45
 290 Thr Gly Tyr His Phe Asn Val Met Thr Cys Glu Gly Cys Lys Gly Phe
 291 50 55 60
 292 Phe Arg Arg Ala Met Lys Arg Asn Ala Arg Leu Arg Cys Pro Phe Arg
 293 65 70 75 80
 294 Lys Gly Ala Cys Glu Ile Thr Arg Lys Thr Arg Arg Gln Cys Gln Ala
 295 85 90 95

VERIFICATION SUMMARY

PATENT APPLICATION: **US/09/276,935C**

DATE: 12/12/2002

TIME: 16:44:09

Input Set : **A:\PU3474US 11-02 Seqlist.txt**

Output Set: **N:\CRF4\12122002\I276935C.raw**

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date

Raw Sequence Listing Error Summary

O1PE

ERROR DETECTEDSUGGESTED CORRECTION

SERIAL NUMBER: 09/276,935C

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 Wrapped Nucleic
Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping".

2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 Misaligned Amino
Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 PatentIn 2.0
"bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s). Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7 Skipped Sequences
(OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8 Skipped Sequences
(NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 Use of n's or Xaa's
(NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10 Invalid <213>
Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11 Use of <220>
- Sequence(s) 11, 12, 14 missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

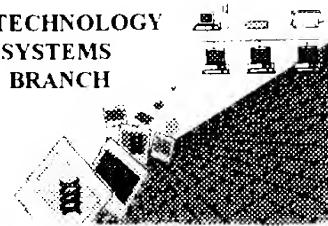
12 PatentIn 2.0
"bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 Misuse of n n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

RECEIVED

DEC 30 2002

TECH CENTER 1600/2900



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

RECEIVED

Application Serial Number: 31276,1350
Source: CDP2

DEC 30 2002

Date Processed by STIC: 12/12/02

TECH CENTER 1600/29

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. **EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>), EFS Submission User Manual - ePAVE)**
2. **U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202**
3. **Hand Carry directly to:**
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202**